Journal of Forestry Research (2011) 22(2): 275–280
DOI 10.1007/s11676-011-0162-z

ORIGINAL PAPER

Comparison of the fraction of olfactory receptor pseudogenes in wolf (*Canis lupus*) with domestic dog (*Canis familiaris*)

ZHANG Hong-hai • WEI Qin-guo • ZHANG Huan-xin • CHEN Lei

Received: 2010-03-26; Accepted: 2010-08-15

Abstract: Olfactory receptors (ORs), the first dedicated molecules with which odorants physically interact to arouse an olfactory sensation, constitute the largest gene family in vertebrates. Dogs and wolves, like many other mammals, have a highly developed capability to detect and identify odorant molecules, even at minimum concentrations. In this study, the olfactory receptor repertoire from domestic dog and its closest relative, the wolf, were sequenced to estimate the fraction of pseudogenes in each subspecies. The fraction of disrupted olfactory receptor genes in dog was 17.78%, whereas, that in wolf was 12.08%. As expected the dog was less dependent on olfaction than the wolf, and the dog had more olfactory receptor pseudogenes. However, the observed difference between the two subspecies was not at the significant level (χ^2 = 1.388, p = 0.239 > 0.05). The values indicated that although domestication might play a role in the reduction of OR genes, it could not be concluded that the living environment provided by domestication lead to a significant reduction of the functional olfactory receptor repertoire. Furthermore, the purpose of domestication may also have influence on the ratio of functional olfactory receptor genes reduction.

Keywords: olfactory receptor; wolf; domestic dog; pseudogene

Introduction

Olfaction, the sense of smell, is an important way used by animals to communicate with environment and their companions. It presents in most animal species and the mechanisms are shared among distantly related taxonomical orders (Strausfeld and Hildebrand 1999). Olfactory receptor (OR) genes were first discovered in *Rattus norvegicus* in 1991 by Buck and Axel (Buck and Axel 1991). Olfactory receptors proteins are expressed on the surface of the cilia of the olfactory sensory neurons lining the neuroepithelium in the nasal cavity and are capable of binding odorants. For discriminating among the millions of possible odorants, the mammalian olfactory system possesses a repertoire of approximately more than 1000 different olfactory receptors. All of the olfactory receptors belong to the G protein-coupled receptor superfamily and are evolutionarily related (Young et al. 2002).

Animal species can be subdivided depending on how critical the olfaction is for their survival (Scott and Fuller 1965; Gilad et al. 2004). It is supposed that in primates such as humans and apes, the senses rather than the olfaction may be more important than smell. This strategy is reflected in some morphological features of primates such as a relatively small nose and large frontal eyes (Gilad et al. 2003). And the reduced dependence on smell is also evident on the molecular level. Therefore, it has been shown that humans have an extremely reduced repertoire compared to the mouse or even to other primates. In humans, about 60% of the olfactory receptors genes carry additional disruptions in their coding region (Glusman et al. 2000), while this fraction is equal to 40% in the chimpanzee and 20% in the mouse (Rouquier et al. 2000). One possible reason for the exceptionally high fraction of pseudogenes in humans is that the conservation of olfactory receptors in humans was relaxed as the special human lifestyle. If the lifestyle is critical factor leading to this exceptionally high fraction of pseudogenes, other species which underwent a similar change in lifestyle should exhibit a similarly changed functional olfactory repertoire.

In order to verify this hypothesis, in the present study the olfactory receptor repertoire from domestic dog and its closest relative, the wolf, were sequenced to estimate the fraction of pseudogenes in each subspecies.

Foundation project: This research was supported by Program for New Century Excellent Talents (NCET-07-0507), National Natural Science Foundation of China (30370218), Natural Science Foundation of Shandong Province (Z2008D01), Project of Science and Technology Development Plan of Shandong Province (2007GG2009011).
The online version is available at http://www.springerlink.com

ZHANG Hong-hai (✉) • WEI Qin-guo • ZHANG Huan-xin • CHEN Lei
College of Life Science, Qufu Normal University, Qufu 273165, Shandong Province, China. E-mail: zhanghonghai67@126.com

Responsible editor: Yu Lei

Materials and methods

Genomic DNA extraction

Blood sample from a female wolf was obtained from the Dalai Lake Nature Reserve in Inner Mongolia of China, and the other blood sample was obtained from a female domestic dog in Shandong province of China. Total genomic DNAs were extracted from the blood following the method of Sambrook and Russell (2001).

Degenerate primer construction

Two sets of degenerate primers were designed on the conserved regions of Human OR genes 1 and 7 (H1 and H7), and each set should amplified approximately 700 bp of the coding region. Primer pair PC1 was designed on human Class I OR genes only, primer pair PC2 on human Class II with the exclusion of subfamily 7E. Both primers were 26 nucleotides long. Sequences are as follows:

PC1-5:5'-CT(GC)CA(CT)(GC)A(AG)CCCA-TGT(AT)(CT)(ACT)(AT)(CT)TT(GCT)CT-3';

PC1-3:5'-GT(CT)(CT)T(GC)A(CT)(AGT)C(ACT)(AG)TA(AG)A(CT)(AG)A(CT)(AG)GGG TT-3';

PC2-5:5'-(CT)T(AGCT)CA(CT)(AT)C(ACT)CC(ACT)ATGTA(CT)TT(CT)TT(GCT)CT-3';

PC2-3:5'-TT(CT)CT(AGCT)A(AG)G(GC)T(AG)TAGAT(AGCT)A(AGCT) (AGT) GG(AG) TT-3'.

The PC1 pair had a total degeneracy of 1.06×10^7, i.e.: many primer combinations were potentially present in the reaction mixture. This value for PC2 is 2.12×10^7. In human, primer pair PC1 amplifies an approximately equal proportion of Class I and Class II OR genes. Primer pair PC2 amplifies exclusively Class II OR genes. To compensate for this effect the final sample was constructed to contain 25% of sequences amplified by primer pair PC1.

Primers were designed on all the OR sequences contained in the HORDE database (Godfrey et al. 2004) and experimentally tested on the mouse and the domestic dog DNAs. For these species, the family composition and the exact fractions of pseudogenes were known (Zhang and Firestein 2002).

Degenerate PCR

PCR was performed in a total volume of 50 µL containing 0.25 mmol of each deoxynucleotide (Promega, Madison WI (9)), 10 pmol each of forward and of reverse primer, 25 mmol MgCl2, 10 mmol Tris pH 8.3, five units of Taq DNA polymerase and 50 ng of genomic DNA. PCR conditions were as follows: 5 min at 95°C followed by 45 cycles of denaturation at 95°C for 40 s, annealing at 47°C for 30 s and elongation at 72°C for 2 min; followed by an additional elongation step at 72°C for 10 min.

Cloning

PCR products were visualized on a 1.8% agarose gel and subsequently cloned AE into chemocompetent E. coli using the pGEM-T Easy TA Cloning Kit (Invitrogen, Carlsbad, CA (5)) following the standard protocol (Version P) with 4 µL of PCR products and a ligation time of 30 min.

Colony PCR

Single bacterial white colonies where picked and suspended in 40 µL of PCR reaction mix and amplified by touch down PCR using the vector primers. Conditions of the touch down PCR were as follows: an initial denaturation step at 94°C for 5 min, 40 cycles of denaturation at 94°C for 30 s, annealing at 57°C for 1 min and elongation at 72°C for 1 min, followed by a terminal elongation step at 72°C for 15 min.

Sequencing

PCR products (5 µL) of positive colony were visualized on a 1.8% agarose gel and the remaining product was purified using the method of ethanol/EDTA/NaAC. Sequencing reactions were performed on PCR products in both directions with a dye-terminator cycle sequencing kit on ABI PRISM 3700 automated DNA sequencer (Applied Biosystems, Foster City, CA (1)).

Sequence analysis

The base was called with the ANALYSIS software, Ver. 3.0 (Applied Biosystems, Foster City, CA (1)). The sequences were visually inspected and edited by SEQMAN, Ver. 5.00 (DNASTAR, Madison, WI (2)). Primer sequences were removed and individual clones with more than 98% identity were joined into one consensus sequence. This cut-off ensured that Taq-generated mutations that might have been sequenced in individual clones were not counted as independent genes. Then we did blast in NCBI. Sequences with no best-hit to known OR genes were removed.

DNAsp4 was used to identify stop codons within the sequences obtained. If there was no interrupted frame, the sequence was considered as a functional gene. If one or more additional interrupted coding region could be found, the sequence was assigned as pseudogene. In some cases, the combining of several clones can create some ambiguous positions. We checked all nucleotide variants of the ambiguous sequences with the same set of rules. If a complete open reading frame existed only resulted from these ambiguous positions and the alternative would result in a stop codon, the sequence was removed from the analysis.

The observed differences in proportions of pseudogenes between wolf and dog were calculated by χ^2 test.

Sample representation

Degenerate primers were chosen to get an overview over the

domestic dog and the wolf OR repertoires. Nevertheless, using degenerate primers might have risk of biased amplification certain genes rather than others. The result from these primers could not represent the familial composition of the OR repertoire. On the other hand, OR families contain different fractions of pseudogenes. Therefore, estimating the OR pseudogene fraction for the entire repertoire could be meaningless. A pilot project was carried out to test the performance of the degenerate primers prior to sampling OR genes from domestic dog and wolf.

A set of degenerate primers were designed on the basis of the H2 (forward primer) and H7 (reverse primer) domains of all human OR genes according to the HORDE database, using the program DEFOG (Fuchs et al. 2002; Matasci 2003). The primer set PC1 was designed on all Class I Ors, while the primer set PC2 was designed on all Class II OR genes. 60 OR genes were amplified from dog genomic DNA with both primer sets. The result showed that primer set PC1 amplifies OR genes from Class I and Class II in equal proportions, and PC2 amplifies Class II OR genes only. It was induced that a sample composed of 25% of OR genes amplified by PC1 and 75% of OR genes amplified by PC2 would reflect the familial composition of OR genes seen in dogs. It was confirmed that the estimation was correct through the analysis of a sample of 100 OR genes using the 25% PC1+75% PC2 rule for domestic dog.

In an analysis of the olfactory repertoire of the dog, both a degenerate PCR based approach and a data-mining of the dog draft genome approach have been undertaken. Both approaches resulted in a similar estimate of the fraction of pseudogenes present in the OR repertoire (Olender et al. 2004). Therefore, according to the analysis, the sample would reflect the familial composition of the OR repertoire of the corresponding species.

Sampling of OR genes

The required number of at least 100 OR genes were selected from 700 sequenced bacterial colonies for each subspecies. The genes of insufficient sequence quality or read length were removed from our test. And if the similarity of two genes exceeded 98%, they were considered as one gene. At last, 127 OR genes were got for domestic dog and 131 for wolf. 108 sequences of correct length were collected for the wolf, 26 from the amplification with the PC1 primers and 82 from the one with the PC2. For dog the number of sequences for the test amplified with PC1 was 26, whereas, the number for PC2 was 81. Through this selection, it was confirmed that the comparison was significative.

Results and discussions

Estimate the fraction of OR pseudogenes

As known, the OR genes have no introns in their coding frame. Pseudogenes were defined as carrying an in-frame premature stop codon within the coding region. Premature stop codons could be the result of nonsense point mutations or frameshift mutations caused by insertions or deletions. These ORs with a premature stop codon were functional, since the 700 bp region amplified in the study was located between the transmemebrane domains H2 and H7, which were considered to be of critical importance for the functioning of these proteins (Buck and Axel 1991). Because of the OR genes encoded by a coding region with no introns, such stop codons could not be overlooked by certain mechanisms such as alternative splicing. On the other way, assessing the fraction of pseudogenes by scanning only 700 bp resulted in an underestimate of the real fraction of gene disruptions, which were likely to occur over the entire coding region. OR genes were aligned using the ClustalW algorithm and dendrograms were visually inspected (Fig. 1(a)). No clusters composed exclusively of pseudogenes could be detected.

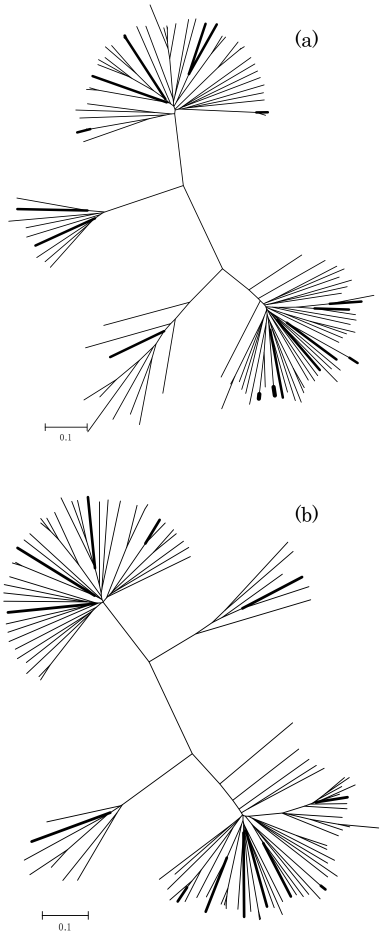

Fig. 1 Unrooted neighbour-joining trees based on nucleotide sequence distances for watchdog (a) and wolf (b). Branch lines: slim lines, functional genes; thick lines, pseudogenes.

The fraction of OR genes could be determined for a set of 107 dog OR genes and 108 wolf OR genes. For dog the fraction of pseudogene is 17.78%, while that for wolf is 12.08%. The observed difference in proportions of pseudogenes is not significant (χ^2 =1.388, p=0.239>0.05).

Potential factors for the fraction of observed pseudogenes

Some other factors rather than biological ones might also influ-

ence the fraction of pseudogenes observed in the two subspecies. As shown in the pilot study, the bias of the degenerate primers towards certain families might significantly influence the fraction of observed pseudogenes. In the study, the results of dogs were compared with the ones obtained by genome-mining analysis of this species (Matarazzo et al. 1998). The comparison showed that the comparison in the study between dogs and wolf is meaningful. And even if there was the deviation, the results could not be significantly affected due to the similar familial composition in the two subspecies. On the other hand, this problem would affect both subspecies alike, resulting in two estimates which are equally biased. Furthermore, the cladograms depicted in Fig. 1(a) and (b), where no major branch exclusively composed of pseudogenes could be detected, provided additional support for the fact that the primers used in this study are unbiased.

Mistakes in the DNA sequence resulted in another possible source of error. While incorrect base calling rarely caused stop codons (<5% of random codon mutations would result in a stop codon), erroneous insertion or deletion within the coding region resulted in an inferred protein sequence, which usually could not bear any similitude with the original one and was extremely likely to be truncated. For such base calling problems depending on the quality of the reads of the DNA sequences amplified and minimizing the difference between the two subspecies, all laboratory operations were performed in parallel. Furthermore, we used only good quality sequences with reads from both strands.

In a study on OR genes from pig (Wagner and Lykke 2002), 15 OR genes sequences were detected via amplification of cDNA obtained from mRNA extracted from olfactory epithelium. But none of these OR genes were pseudogenes. We think that OR genes that carry destructive mutation are already pseudogenes at first, and they might have already been switched off by some other means, possibly acting on promoters and enhancers. Mutations can accumulate without any effect. The proteins that originally encoded by these genes are not produced anymore, even dramatic changes like frame shift or non-sense mutations can accumulate in the genome. Additionally, even if mRNA encoding for such non-functional genes is produced, it might be recognized as bad and immediately destroyed by the non-sense mediated RNA decay mechanism (Mombaerts 1999) Such factors may explain why mRNA of OR pseudogenes could not be found in these studies. The same phenomenon can be observed in the mouse. None pseudogene was found in the 55 OR genes detected in mouse from olfactory epithelium, while the actual value for the fraction of pseudogenes in mouse was 20% (Zhang and Firestein 2002).

Fraction of pseudogenes

According to the result of previous studies, 60% or more of the human OR genes carry one or more disruptions (Glusman et al. 2001), on the other way, the overall number of OR genes is similar in all mammals. It is a hypothesis that the loss of functional OR genes is owing to a relaxation of evolutionary constraints on olfaction resulting from the human lifestyle. For investigating this hypothesis further, it is to inferred whether domesticated species have a higher proportion of pseudogenes in their OR repertoire than their closest wild living relatives. The proportion of OR pseudogenes was estimated to find the key to this question in the repertoires of a domestic dog and a wolf. It was found that the fraction of pseudogenes defined as interruption of the inferred protein sequence is 17.78% in the dog, while that in the wolf is 12.08%, which is 5.70% lower than its closest domestic relative. However, this difference is not statistically significant. If these estimates represent the real proportions of pseudogenes in each subspecies, it would require a much larger sample (roughly 300 genes from each subspecies) in order to have enough power to achieve statistical significance at the 5% level. In this study, we does not consider genes with interrupted coding region are not expressed, correctly folded or transported to the cell surface to perform their function. These are also non-functional genes, which are much more difficult to identify. The fractions presented here, as well as in other studies, are thus lower estimates.

OR familial composition in dog and wolf

From the results of the studies (Table 1 and 2), the difference in the relative sizes of OR families between the two subspecies is relatively limited. Therefore, it is concluded that the comparison is meaningful.

Table 1. OR familial composition and pseudogene fraction in domestic dog

Family	OR No.	Family (%)	Pseudogenes No.	Pseudogenes fraction
1	2	1.87%	0	0.00%
2	9	8.41%	0	0.00%
3	4	3.74%	0	0.00%
4	8	7.48%	0	0.00%
5	8	7.48%	1	12.50%
6	4	3.74%	1	25.00%
7	6	5.61%	1	16.67%
8	14	13.08%	2	14.29%
9	3	2.80%	1	33.33%
10	13	12.15%	3	20.08%
11	5	4.67%	4	80.00%
13	7	6.54%	3	42.86%
51	5	4.67%	1	20.00%
52	10	9.35%	2	20.00%
56	1	0.93%	0	0.00%
others	8	7.48%	0	0.00%
total	107	100%	19	17.78%

Comparison with other mammals

The OR familial composition is similar across mammals, rodents and carnivores, so the figures can be compared with the values obtained in similar studies. Especially, since we used the same set of primers as in Gilad et al. (2003), we will compare the results of this study with these studies. From the comparison, the fraction of pseudogenes in dog and wolf are lower than that in

other mammals (Fig. 2). These data might supply some molecular basis of canine prominent chemosensory capacities.

Table 2. OR familial composition and pseudogene fraction in wolf

Family	OR No.	Family (%)	Pseudogenes No.	Pseudogenes fraction
1	2	1.85%	0	0.00%
2	11	10.19%	1	9.09%
3	1	0.93%	0	0.00%
4	4	3.70%	0	0.00%
5	6	5.56%	1	16.67%
6	0	0.00%	0	0.00%
7	16	14.81%	0	0.00%
8	13	12.04%	0	0.00%
9	5	4.63%	0	0.00%
10	10	9.26%	1	10.00%
11	4	3.70%	2	50.00%
13	7	6.49%	3	42.86%
51	6	5.56%	0	0.00%
52	6	5.56%	1	16.67%
56	1	0.93%	1	100.00%
others	16	14.81%	3	18.75%
total	108	100%	13	12.08%

Domestication, a possible reason for the functional ORs genes reduction

The difference between dogs and wolf, although not significant, goes in the direction expected if domestication was the causes. Dogs appear to have indeed a smaller functional OR repertoire than wolf. In a domestic environment, dogs rely on humans especially for activities, while olfaction plays a critical role in the wild subspecies, such as food, friend location and other social interactions. In such a case, OR genes are lost by mutation and random drift. For example, random mutations can accumulate to lead to the inactivation of the gene, and the mutated gene can increase its frequency in the population. However, mutations that do not affect the fitness usually require long time to spread within the population and the reach fixation. The time elapsed since domestication (15 000–17 000 years) (Savolainen et al. 2002) can hardly account for such an effect.

On the other hand, the loss of OR could be the result of an active selective process. In contrast to neutral traits, traits under selection get fixed rapidly, as few as 10 generations can suffice, and domestication provides very strong selection. An example of such phenomena is the connection between agouti coat color and tameness observed in several mammals (Price 2002).

The changes in nasal airways might be an adaptive change as a result of short time domestication. It is the strengthening hypothesis that olfaction, which is especially apparent between pig and wild boar (Olender et al. 2004), might have been under some sort of selection during domestication (Price 2002).

However, there was no significant difference to conclude that the living environment provided by domestication played a major role in the reduction of the functional olfactory receptor repertoire. Therefore, the purpose of domestication may also play a key role on dogs' olfaction. Olfaction also plays a key role in the behavior of dogs and in its social life. The reduction of dogs' OR genes is not apparent.

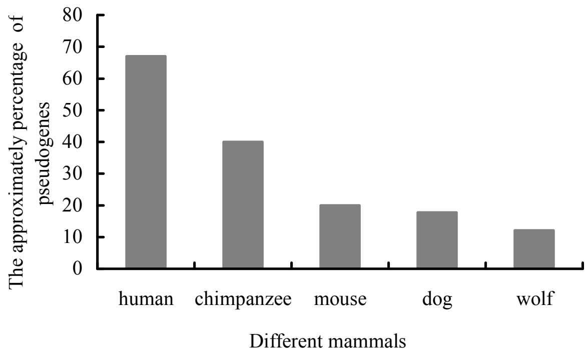

Fig. 2 Comparison of the fraction of pseudogenes in different mammals (Godfrey et al. 2004; Olender et al. 2004; Rouquier et al. 2000)

Conclusion

According to the present results and other researches carried out previously, it is inferred that domestication might lead to the reduction of OR genes at a certain extent, while the purpose of domestication also plays an important role on the selection of dogs' olfaction. During the course of evolution, there were various factors influencing the change of olfaction, and therefore, much more studies have to be carried out to explain the evolutional meaning from wolf to dog.

References

Buck L, Axel R. 1991. A novel multigene family may encode odorantreceptors: a molecular basis for odor recognition. *Cell*, **65**(1): 175–187.

Fuchs T, Malecova B, Linhart C, Sharan R, Khen M, Herwig R, Shmulevich D, Elkon R, Steinfath M, O'Brien JK, Radelof U, Lehrach H, Lancet D, Shamir R. 2002. DEFOG: a practical scheme for deciphering families of genes. *Genomics*, **80**(3): 295–302.

Gilad Y, Man O, Pääbo S, Lancet D. 2003. Human specific loss of olfactory receptor genes. *Proc Natl Acad Sci*, **100**(6): 3324–3327.

Gilad Y, Przeworski M, Lancet D. 2004. Loss of olfactory receptor genes coincides with the acquisition of full trichromatic vision in primates. *PLoS Biology*, **2**(1): E5.

Glusman G, Bahar A, Sharon D, Pilpel Y, White J, Lancet D. 2000. The olfactory receptor gene superfamily: data mining, classification, and nomenclature. *Mamm Genome*, **11**(11): 1016–1023.

Glusman G, Yanai I, Rubin I, Lancet D. 2001. The complete human olfactory subgenome, *Genome Res*, **11**(5): 685–702.

Godfrey PA, Malnic B, Buck LB. 2004. The mouse olfactory receptor gene family. *Proc Natl Acad Sci*, **101**: 2156–2161.

Lander ES, Linton LM, Birren B et al. 2001. Initial sequencing and analysis of the human genome. *Nature*, **409**(6822): 860–921.

Matarazzo V, Tirard A, Renucci M, Belaïch A, Clément JL. 1998. Isolation of

putative olfactory receptor sequences from pig nasal epithelium. *Neuroscience Letters*, **249**(2–3): 87–90.

Matasci N. 2003. Did Domestication Reduce the Fraction of Functional Olfactory Receptors in Pig? Max Planck Institut, Zurich.

Mombaerts P. 1999. Molecular biology of odorant receptors in vertebrates. *Annual Review of Neuroscience*, **22**: 487–509.

Olender T, Fuchs T, Linhart C, Shamir R, Adams M, Kalush F, Khen M, Lancet D. 2004. The canine olfactory subgenome. *Genomics*, **83**(3): 361–372.

Price EO. 2002. Approaches to the study of domestication. Animal Domestication and Behaviour. E. O. Price. Wallingford, Oxon, UK, CABI Publishing: 13–20.

Price EO. 2002. Behavioural development in captive animals. Animal domestication and behavior. E. O. Price. Wallingford, Oxon, UK, CABI Publishing, pp.161–181.

Rouquier S, Blancher A, Giorgi D. 2002. The olfactory receptor gene repertoire in primates and mouse: evidence for reduction of the functional fraction in primates. *Proc Natl Acad Sci*, **97**: 2870–2874.

Sambrook J, Russell DW. 2001. Molecular Cloning: A Laboratory Manual. NY: Cold Spring Harbor Press, Cold Spring Harbor.

Savolainen P, Zhang YP, Luo J, Lundeberg J, Leitner T. 2002. Genetic Evidence for an East Asian Origin of Domestic Dogs. *Science*, **298**(22): 1610–1613.

Scott JP, Fuller JL. 1965. Genetics and the Social Behavior of the Dog. Chicago: Univ Chicago Press.

Strausfeld NJ, Hildebrand JG. 1999. Olfactory systems: common design, uncommon origins? Current Opinion in Neurobiology, **9**(5): 634–639.

Wagner E, Lykke AJ. 2002. mRNA surveillance: the perfect persist. *J Cell Sci*, **115**(Pt 15): 3033–3038.

Young JM, Trask BJ. 2002. The sense of smell: genomics of vertebrate odorant receptors. *Hum Mol Genet*, **11**(10):1153–1160.

Zhang X, Firestein S. 2002. The olfactory receptor gene superfamily of the mouse. *Nature Neuroscience*, **5**(2): 124–133.